



OIPE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/983,000

DATE: 11/02/2001  
TIME: 15:30:46

Input Set : A:\es.txt  
Output Set: N:\CRF3\11022001\I983000.raw

**Does Not Comply  
Corrected Diskette Needed**

*Errors on pp. 13, 5, + on sequences as  
noted below.*

3 <110> APPLICANT: AGY Therapeutics  
4 Melcher, Thorsten  
5 Mueller, Sabine  
6 Chin, Daniel  
8 <120> TITLE OF INVENTION: USE OF PROTEIN BIOMOLECULAR TARGETS IN THE TREATMENT AND  
VISUALIZATION OF  
9 BRAIN TUMORS  
11 <130> FILE REFERENCE: 263/180 -- PEagleman -- AGY  
C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/983,000  
C--> 13 <141> CURRENT FILING DATE: 2001-10-17  
13 <160> NUMBER OF SEQ ID NOS: 35  
15 <170> SOFTWARE: PatentIn version 3.1

#### ERRORED SEQUENCES

6210 <210> SEQ ID NO: 23  
6211 <211> LENGTH: 3143  
6212 <212> TYPE: DNA  
6213 <213> ORGANISM: Homo sapiens  
6215 <220> FEATURE:  
W--> 6216 <221> NAME/KEY: REFERENCE  
6217 <222> LOCATION: (1)..(3143)  
6218 <223> OTHER INFORMATION: Crosby, A.H., Edwards, S.J., Murray, J.C. and Dixon, M.J.  
TITLE G  
6219 enomic organization of the human osteopontin gene: exclusion of t  
6220 he locus from a causative role in the pathogenesis of dentinogene  
6221 sis imperfecta type II JOURNAL Genomics 27 (1), 155-160 (1995)  
6224 <220> FEATURE:  
6225 <221> NAME/KEY: Gene  
6226 <222> LOCATION: (1)..(3143)  
6227 <223> OTHER INFORMATION: Osteopontin  
6230 <220> FEATURE:  
6231 <221> NAME/KEY: CDS  
6232 <222> LOCATION: (374)..(427)  
6233 <223> OTHER INFORMATION:  
6236 <220> FEATURE:  
6237 <221> NAME/KEY: CDS  
6238 <222> LOCATION: (537)..(575)  
6239 <223> OTHER INFORMATION:  
6242 <220> FEATURE:  
6243 <221> NAME/KEY: CDS  
  
6244 <221> NAME/KEY: CDS  
6250 <222> LOCATION: (1232)..(1273)  
6251 <223> OTHER INFORMATION:

*The type of errors shown exist throughout  
the Sequence Listing. Please check subsequent  
sequences for similar errors.  
for Seq. 6, 7, 11, 13, 15, 17, 19, 21*

*Please see Sec. 1.823 of the Rules.  
This information belongs in the  
<3007 to <309> fields*

*Please see example  
proper  
attached.*

6254 <220> FEATURE:

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/983,000

DATE: 11/02/2001

TIME: 15:30:47

Input Set : A:\es.txt

Output Set: N:\CRF3\11022001\I983000.raw

6255 &lt;221&gt; NAME/KEY: CDS

6256 &lt;222&gt; LOCATION: (1597)..(1920)

6257 &lt;223&gt; OTHER INFORMATION:

6260 &lt;220&gt; FEATURE:

6261 &lt;221&gt; NAME/KEY: CDS

6262 &lt;222&gt; LOCATION: (2305)..(2709)

6263 &lt;223&gt; OTHER INFORMATION:

*discrepancy:*E--> 6266 <400> SEQUENCE: 33 → *<2107 is listed as 23*

```

6267 ggggaagtgt gggagcaggt gggctgggca gtggcagaaa cctgatgaca caatctcgcc      60
6269 gcctccctgt gttggtggag gatgtctgca gcagcattta aattctggga gggcttggtt      120
6271 gtcagcagca gcaggaggag gcagagacag catcgtcggg accagactcg tctcaggcca      180
6273 gttgcagcct tctcagccaa acgccgacca aggtacagct tcagtttgct actgggttgt      240
6275 gcattcaqct qaatttcatt ggggaagtcca aattctaagg aaaaaaatgt ggtagtataa      300
6277 aaaggatatca ctgttgtaac ctatgaagat gtcagctatt cctttgaaat attttgcagg      360
6279 aaaactcact acc atg aga att gca gtg att tgc ttt tgc etc cta ggc      409
6280                               Met Arg Ile Ala Val Ile Cys Phe Cys Leu Leu Gly
6281                               1                               5                               10
6283 atc acc tgt gcc ata cca gtgagtacag ttgcatttta aagaaaattc      457
6284 Ile Thr Cys Ala Ile Pro
6285                               15
6287 ctgaaaataa ctgaattgtg tgcttccatg tgctaggagg acattcttgt aatcttttct      517
6289 catcttttct gtttctaag gtt aaa cag gct gat tct gga agt tct gag gaa      569
6290                               Val Lys Gln Ala Asp Ser Gly Ser Ser Glu Glu
6291                               20                               25
6293 aag cag gtaagcatct tttatgtttt tatatagtta aatcatttac tcaattatgg      625
6294 Lys Gln
6295 30
6297 cgagaggtgc aagaaaacgta tttgctgcga tcaaatgagt tcatatttgt aaagcaattt      685
6299 gaaagagtgc ctagcccaca gtaagtgcta cataagagtt tgtaaataa atctgcaaaa      745
6301 aaaaaaaaaa ttacaaaaag gtacctaaag gtcggggtga ctatatgctt ccatcaagac      805
6303 tagtgaagaa tggttgtttt ttccattcat cctacattt ctttttttaa taatgataaa      865
6305 catgcaactt tttttag ctt tac aac aaa tac cca gat gct gtg gcc aca      916
6306                               Leu Tyr Asn Lys Tyr Pro Asp Ala Val Ala Thr
6307                               35                               40
6309 tgg cta aac cct gac cca tct caq aag cag aat etc cta gcc cca cag      964
6310 Trp Leu Asn Pro Asp Pro Ser Gln Lys Gln Asn Leu Leu Ala Pro Gln
6311                               45                               50                               55
6313 gtatttttaa acttctcata attaaactac agtgatgaaa qatagccaca ctcaggccat      1024
6315 ttgggctgct cagatgaatc ctgccctgcc tgctggcaaa catgtgctta ggacattgac      1084
6317 tgatctgcca tgttgcttc tctctgtgtt aagccatcca cagatgaggc tgaaaaataa      1144
6319 aaactgcttt ggattaaaaa ggttaacttt tgaataaaaa agctaggcat gtgtgatgag      1204
6321 cactaacacg tgccattcct tcttcag aat gct gtg tcc tct gaa gaa acc aat      1258
6322                               Asn Ala Val Ser Ser Glu Glu Thr Asn
6323                               60                               65
6325 gac ttt aaa caa gag gtaagttctc attttcaatc agaaqcccat catgccttga      1313
6326 Asp Phe Lys Gln Glu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/983,000

DATE: 11/02/2001

TIME: 15:30:47

Input Set : A:\es.txt

Output Set: N:\CRF3\11022001\I983000.raw

```

6333 aataactaggt ttctctggat aacctaaaaag ccatggtatg tactgtgaat gcaaagattc 1493
6335 tgaaactaaa taaaaagaaa gatagtaaaa gactaatgtg ctataaaggc taagggaata 1553
6337 taaaaaccca tatattaatt ttcccggcca tcttaatttt cag acc ctt cca agt 1608
6338                                     Thr Leu Pro Ser
6339                                     75
6341 aag tcc aac gaa agc cat gac cac atg gat gat atg gat gat gaa gat 1656
6342 Lys Ser Asn Glu Ser His Asp His Met Asp Asp Met Asp Asp Glu Asp
6343                                     80                                     85                                     90
6345 gat gat gac cat gtg gac agc cag gac tcc att gac tcg aac gac tct 1704
6346 Asp Asp Asp His Val Asp Ser Gln Asp Ser Ile Asp Ser Asn Asp Ser
6347                                     95                                     100                                     105
6349 gat gat gta gat gac act gat gat tct cac cag tct gat gag tct cac 1752
6350 Asp Asp Val Asp Asp Thr Asp Asp Ser His Gln Ser Asp Glu Ser His
6351                                     110                                     115                                     120
6353 cat tct gat gaa tct gat gaa ctg gtc act gat ttt ccc acg gac ctg 1800
6354 His Ser Asp Glu Ser Asp Glu Leu Val Thr Asp Phe Pro Thr Asp Leu
6355 125                                     130                                     135                                     140
6357 cca gca acc gaa gtt ttc act cca gtt gtc ccc aca gta gac aca tat 1848
6358 Pro Ala Thr Glu Val Phe Thr Pro Val Val Pro Thr Val Asp Thr Tyr
6359                                     145                                     150                                     155
6361 gat ggc cga ggt gat agt gtg gtt tat gga ctg agg tca aaa tct aag 1896
6362 Asp Gly Arg Gly Asp Ser Val Val Tyr Gly Leu Arg Ser Lys Ser Lys
6363                                     160                                     165                                     170
6365 aag ttt cgc aga cct gac atc cag gtaaatcctt taacagacac acctgatggt 1950
6366 Lys Phe Arg Arg Pro Asp Ile Gln
6367                                     175                                     180
6369 tctgactagc gctcaagtct aggaaaccac agtttgcata ttcattcatt cattcatcca 2010
6371 ttcattcatt cattcagcaa gaattcattc atattctact ttatgacat tgaatacaaa 2070
6373 tctttttctg cttggcggtt tttgtaagtc tacataattt ctctctagat ttgattctca 2130
6375 aacacaattc tactttttga aatcctggat caaagtaaca tgctagtatt atttcagcca 2190
6377 gatttagaca atttttagta taagatgacc taaaagctag agagtggaaa aggattacca 2250
6379 tattcccatc cctagcgtt catataatta ttcttcattt gtgccgtgat tcag tac 2307
6380                                     Tyr
6383 cct gat gct aca gac gag gac atc acc tca cac atg gaa agc gag gag 2355
6384 Pro Asp Ala Thr Asp Glu Asp Ile Thr Ser His Met Glu Ser Glu Glu
6385                                     185                                     190                                     195
6387 ttg aat ggt qca tac aag gcc atc ccc gtt gcc cag gac ctg aac gcg 2403
6388 Leu Asn Gly Ala Tyr Lys Ala Ile Pro Val Ala Gln Asp Leu Asn Ala
6389                                     200                                     205                                     210
6391 cct tct gat tgg gac agc cgt ggg aag gac agt tat gaa acg agt cag 2451
6392 Pro Ser Asp Trp Asp Ser Arg Gly Lys Asp Ser Tyr Glu Thr Ser Gln
6393                                     215                                     220                                     225
6395 ctg gat gac cag agt gct gaa acc cac agc cac aag cag tcc aga tta 2499
6396 Leu Asp Asp Gln Ser Ala Glu Thr His Ser His Lys Gln Ser Arg Leu
6397 230                                     235                                     240                                     245
6399 tat aaq cgq aaa gcc aat gat gag agc aat qaq cat tcc gat gtg att 2547

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/983,000

DATE: 11/02/2001

TIME: 15:30:47

Input Set : A:\es.txt

Output Set: N:\CRF3\11022001\I983000.raw

```

6404 Asp Ser Gln Glu Leu Ser Lys Val Ser Arg Glu Phe His Ser His Glu
6405                265                270                275
6407 ttt cac agc cat gaa gat atg ctg gtt gta gac ccc aaa agt aag gaa      2643
6408 Phe His Ser His Glu Asp Met Leu Val Val Asp Pro Lys Ser Lys Glu
6409                280                285                290
6411 gaa gat aaa cac ctg aaa ttt cgt att tct cat gaa tta gat agt gca      2691
6412 Glu Asp Lys His Leu Lys Phe Arg Ile Ser His Glu Leu Asp Ser Ala
6413                295                300                305
6415 tct tct gag gtc aat taa aaggagaaaa aatacaattt ctcactttgc      2739
6416 Ser Ser Glu Val Asn
6417 310
6419 atttagtcaa aagaaaaaat gctttatagc aaaatgaaag agaacatgaa atgctttcttt      2799
6421 ctcaqtttat tgggtgaatg tgtatctatt tgagtcgga aataactaat gtgtttgata      2859
6423 attagtttag tttgtggctt catggaaact ccctgtaaac aaaagcttca gggttatgtc      2919
6425 tatgttcatt ctatagaaga aatgcaaact atcactgtat tttaatatatt gttattctct      2979
6427 catgaataga aatttatgta gaagcaaaca aaatactttt acccacttaa aaagagaata      3039
6429 taacatttta tgtcactata atcttttggt ttttaagtta gtgtatatatt tgttgtgatt      3099
6431 atcttttggt gtgtgaataa atcttttatc ttgaatgtaa taag      3143
6569 <210> SEQ ID NO: 25
6570 <211> LENGTH: 259
6571 <212> TYPE: PRT
6572 <213> ORGANISM: Homo sapiens
6574 <220> FEATURE:
6575 <221> NAME/KEY: Gene
6576 <222> LOCATION: (1)..(259)
6577 <223> OTHER INFORMATION: Carbonic Anhydrase domain of human carbonic anhydrase III
E-> 6580 <400> SEQUENCE: 41 → discrepancy!
6582 Ala Lys Glu Trp Gly Tyr Ala Ser His Asn Gly Pro Asp His Trp His
6583 1                5                10                15
6586 Glu Leu Phe Pro Asn Ala Lys Gly Glu Asn Gln Ser Pro Ile Glu Leu
6587                20                25                30
6590 His Thr Lys Asp Ile Arg His Asp Pro Ser Leu Gln Pro Trp Ser Val
6591                35                40                45
6594 Ser Tyr Asp Gly Gly Ser Ala Lys Thr Ile Leu Asn Asn Gly Lys Thr
6595                50                55                60
6598 Cys Arg Val Val Phe Asp Asp Thr Tyr Asp Arg Ser Met Leu Arg Gly
6599 65                70                75                80
6602 Gly Pro Leu Pro Gly Pro Tyr Arg Leu Arg Gln Phe His Leu His Trp
6603                85                90                95
6606 Gly Ser Ser Asp Asp His Gly Ser Glu His Thr Val Asp Gly Val Lys
6607                100               105               110
6610 Tyr Ala Ala Glu Leu His Leu Val His Trp Asn Pro Lys Tyr Asn Thr
6611                115               120               125
6614 Phe Lys Glu Ala Leu Lys Gln Arg Asp Gly Ile Ala Val Ile Gly Ile
6615                130               135               140
6618 Phe Leu Lys Ile Gly His Glu Asn Gly Glu Phe Gln Ile Phe Leu Asp

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/983,000

DATE: 11/02/2001

TIME: 15:30:47

Input Set : A:\es.txt

Output Set: N:\CRF3\11022001\I983000.raw

```

6626 Phe Asp Pro Ser Cys Leu Phe Pro Ala Cys Arg Asp Tyr Trp Thr Tyr
6627           180           185           190
6630 Gln Gly Ser Phe Thr Thr Pro Pro Cys Glu Glu Cys Ile Val Trp Leu
6631           195           200           205
6634 Leu Leu Lys Glu Pro Met Thr Val Ser Ser Asp Gln Met Ala Lys Leu
6635           210           215           220
6638 Arg Ser Leu Leu Ser Ser Ala Glu Asn Glu Pro Pro Val Pro Leu Val
6639 225           230           235           240
6642 Ser Asn Trp Arg Pro Pro Gln Pro Ile Asn Asn Arg Val Val Arg Ala
6643           245           250           255
6646 Ser Phe Lys

```

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/983,000

DATE: 11/02/2001

TIME: 15:30:48

Input Set : A:\es.txt

Output Set: N:\CRF3\11022001\I983000.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No  
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:1900 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:6  
L:1915 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:6  
L:2762 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:7  
L:2768 M:259 W: Allowed number of lines exceeded, <223> Other Information:  
L:3408 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:11  
L:4386 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:13  
L:4392 M:259 W: Allowed number of lines exceeded, <223> Other Information:  
L:4786 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:15  
L:5392 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:17  
L:5644 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:19  
L:5650 M:259 W: Allowed number of lines exceeded, <223> Other Information:  
L:6030 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:21  
L:6036 M:259 W: Allowed number of lines exceeded, <223> Other Information:  
L:6216 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:23  
L:6266 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:23 differs:33  
L:6580 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:25 differs:41

<110> Smith, John; Smithgene Inc.

<120> Example of a Sequence Listing

<130> 01-00001

<140> PCT/EP98/00001  
<141> 1998-12-31

<150> US 08/999,999  
<151> 1997-10-15

<160> 4

<170> PatentIn version 2.0

<210> 1  
<211> 389  
<212> DNA  
<213> Paramecium sp

<220>  
<221> CDS  
<222> (279)...(389)

<300>  
<301> Doe, Richard  
<302> Isolation and Characterization of a Gene Encoding a  
Protease from Paramecium sp  
<303> Journal of Genes  
<304> 1  
<305> 4  
<306> 1-7  
<307> 1988-06-31  
<308> 123456  
<309> 1988-06-31

<400> 1  
agctgtagtc attcctgtgt cctcttctct ctgggtctct caccctgcta atcagatctc 60  
aggagagatg tcttgacct cctctgcctt tgcagcttca caggcagga ggcaggcagc 120  
tgatgtggca attgctggca gtgccacagg ctttcagcc aggcctaggg tgggttcgcg 180  
cgcggcgcgg cggccctct cgcgtctctc tgcgcctct ctctcgtct cctctcgtct 240



Appendix 3, page 2

ggacctgatt aggtgagcag gaggagggggg cagtttagc atg gtt tca atg ttc agc 296  
Met Val Ser Met Phe Ser  
1

ttg tct ttc aaa tgg cct gga ttt tgt ttg ttt gtt tgt ttg ttc caa 344  
Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu Phe Val Cys Leu Phe Gln  
10 15 20

tgt ccc aaa gtc ctc ccc tgt cac tca tca ctg cag ccg aat ctt 389  
Cys Pro Lys Val Leu Pro Cys His Ser Ser Leu Gln Pro Asn Leu  
25 30 35

<210> 2  
<211> 37  
<212> PRT  
<213> Paramecium sp.

<400> 2  
Met Val Ser Met Phe Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu  
1 5 10 15

Phe Val Cys Leu Phe Gln Cys Pro Lys Val Leu Pro Cys His Ser Ser  
20 25 30

Leu Gln Pro Asn Leu  
35

<210> 3  
<211> 11  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Designed peptide based on size and polarity to act as a linker between the alpha and beta chains of Protein XYZ

<400> 3  
Met Val Asn Leu Glu Pro Met His Thr Glu Ile  
1 5 10

<210> 4  
<400> 4  
000

[Annex VIII follows]

identifiers and their accompanying information as shown in the following table. The numeric identifier shall be used only in the "Sequence Listing." The order and presentation of the items of information in the "Sequence Listing" shall conform to the arrangement given below. Each item of information shall begin on a new line and shall begin with the numeric identifier enclosed in angle brackets as shown. The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "O" is optional. Numeric identifiers <110> through <170> shall only be set forth at the beginning of the "Sequence Listing." The following table illustrates the numeric identifiers.

Numeric Identifier	Definition	Comments and Format	Mandatory (M) or Optional (O)
<110>	Applicant	Preferably max. of 10 names; one name per line; preferable format: Surname, Other Names and/or Initials	M
<120>	Title of Invention		M
<130>	File Reference	Personal file reference	M, when filed prior to assignment of appl. number
<140>	Current Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if available
<141>	Current Filing Date	Specify as: yyyy-mm-dd	M, if available
<150>	Prior Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if applicable include priority documents under 35 USC 119 and 120
<151>	Prior Application Filing Date	Specify as: yyyy-mm-dd	M, if applicable
<160>	Number of SEQ ID NOS	Count includes total number of SEQ ID NOS	M
<170>	Software	Name of software used to create the Sequence Listing	O
<210>	SEQ ID NO:##	Response shall be an integer representing the SEQ ID NO	M

<212>	Type	Whether presented sequence molecule is DNA, RNA, or PRT (protein). If a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/RNA molecule shall be further described in the <220> to <223> feature section.	M
<213>	Organism	Scientific name, i.e. Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section.	M
<220>	Feature	Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence.	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA
<221>	Name/Key	Provide appropriate identifier for feature, preferably from WIPO Standard ST.25 (1998), Appendix 2, Tables 5 and 6	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence
		first and last bases/amino acids	usual L-amino acid or modified

		in feature	base was used in a sequence
<223>	Other Information	Other relevant information; four lines maximum	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<300>	Publication Information	Leave blank after <300>	0
<301>	Authors	Preferably max of ten named authors of publication; specify one name per line; preferable format: Surname, Other Names and/or Initials	0
<302>	Title		0
<303>	Journal		0
<304>	Volume		0
<305>	Issue		0
<306>	Pages		0
<307>	Date	Journal date on which data published; specify as yyyy-mm-dd, MMM-yyyy or Season-yyyy	0
<308>	Database Accession Number	Accession number assigned by database including database name	0
<309>	Database Entry Date	Date of entry in database; specify as yyyy-mm-dd or MMM-yyyy	0

example: 05  
07/999,999

<311>	Patent Filing Date	Document filing date, for patent-type citations only; specify as yyyy-mm-dd	0
<312>	Publication Date	Document publication date, for patent-type citations only; specify as yyyy-mm-dd	0
<313>	Relevant Residues	FROM (position) TO (position)	0
<400>	Sequence	SEQ ID NO should follow the numeric identifier and should appear on the line preceding the actual sequence	0

5. Section 1.024 is revised to read as follows:

1.024 Form and format for nucleotide and/or amino acid sequence submissions in computer readable form.

(a) The computer readable form required by 1.021(c) shall meet the following specifications:

(1) The computer readable form shall contain a single "Sequence Listing" as either a diskette, series of diskettes, or other permissible media outlined in paragraph (c) of this section.

(2) The "Sequence Listing" in paragraph (a) (1) of this section shall be submitted in American Standard Code for Information Interchange (ASCII) text. No other formats shall be allowed.

(3) The computer readable form may be created by any means, such as word processors, nucleotide/amino acid sequence editors or other custom computer programs; however, it shall conform to all specifications detailed in this section.

(4) File compression is acceptable when using diskette media, so long as the compressed file is in a self-extracting format that will decompress on one of the systems described in paragraph (b) of this section.

(5) Page numbering shall not appear within the computer readable form version of the "Sequence Listing" file.

(6) All computer readable forms shall have a label permanently affixed thereto on which has been hand-printed or typed: the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form, the operating system used, a reference number, and an application serial number and filing date, if known.

(2) Operating System MS-DOS, Unix or Macintosh.